

**Data Technician**

|  |
| --- |
|  |

|  |
| --- |
| Name: |
| Course Date: |
|  |

**Table of contents**

[Day 1: Task 1 3](#_Toc1514505634)

[Day 1: Task 2 3](#_Toc2125927931)

[Day 1: Task 3 4](#_Toc1965065538)

[Day 2: Task 1 5](#_Toc2141651249)

[Day 3: Task 1 9](#_Toc1229522580)

[Day 3: Task 2 10](#_Toc508692763)

[Day 3: Task 3 11](#_Toc1233463339)

[Day 4: Task 1 12](#_Toc1556426903)

[Day 4: Task 2 13](#_Toc214156810)

[1. Scenario Background 13](#_Toc248266112)

[2. Data Laws and Regulations 14](#_Toc1387014804)

[3. Azure Service Recommendations 14](#_Toc1952198484)

[4. Data Types and Data Modelling 14](#_Toc1792190821)

[5. Data Storage Formats and Structures in Azure 14](#_Toc385598743)

[6. Additional Considerations 15](#_Toc561077662)

[Submission Guidelines: 15](#_Toc1001523541)

[Course Notes 17](#_Toc977988415)

[Additional Information 17](#_Toc1081373283)

# Day 1: Task 1

Please research and complete the below questions relating to key concepts of cloud.

Be prepared to discuss the below in the group following this task.

|  |  |
| --- | --- |
| What can cloud computing do for us in the real-world? | Cloud computing is one of the biggest technological shifts in the last two decades — it lets individuals, businesses, and governments access computing power, storage, and services over the internet instead of relying solely on local hardware. In the real world, it enables a lot of practical, everyday benefits.  Here are some key examples:  1. Accessible Data and Applications Anywhere Example: Google Drive, Dropbox, and Microsoft 365 let users access documents and collaborate in real time from any device, anywhere. Impact: Remote work, online learning, and global collaboration are possible because of cloud-hosted platforms.  2. Scalable Storage and Computing Power Example: A video streaming company (like Netflix) can scale up computing resources to handle millions of users during peak hours — and scale down when traffic drops. Impact: Companies only pay for what they use, avoiding the cost of buying and maintaining large data centers.  3. Cost Savings and Efficiency Example: Small startups can host their apps on Amazon Web Services (AWS), Microsoft Azure, or Google Cloud instead of buying expensive servers. Impact: Reduces upfront costs and allows even small teams to run powerful applications.  4. AI and Big Data Capabilities - Example: Cloud-based machine learning platforms (like AWS SageMaker or Google AI) help businesses analyse massive datasets or deploy AI models without owning supercomputers. Impact: Enables innovations such as fraud detection, medical image analysis, and predictive maintenance.  5. Data Backup and Disaster Recovery Example: Hospitals or banks use cloud backup systems to automatically store copies of critical data off-site. Impact: Protects organisations from data loss due to hardware failure, cyberattacks, or natural disasters.  6. Powering Everyday Services Example: Streaming (Netflix), gaming (Xbox Cloud Gaming), navigation (Google Maps), and social media (Instagram) all rely on cloud servers. Impact: Most modern digital experiences wouldn’t function without the cloud.  7. Faster Software Development and Deployment Example: Developers use cloud-based tools like GitHub Actions or Jenkins to test and deploy apps automatically. Impact: Accelerates innovation and supports continuous integration/delivery (CI/CD).  8. Environmental and Resource Efficiency Example: Cloud data centers' are often more energy-efficient than scattered on-premise servers. Impact: Consolidation and optimisation can reduce global energy consumption and carbon footprints. |
| How can it benefit a business? | Cloud computing can transform how businesses operate, no matter their size or industry. It offers flexibility, cost-efficiency, and tools that help companies innovate faster and stay competitive.  **1. Reduces Costs**   * **No need for physical servers:** Businesses don’t have to buy, maintain, or upgrade expensive hardware. * **Pay-as-you-go model:** You only pay for what you use (storage, computing power, etc.). * **Lower energy and maintenance costs:** Cloud providers handle data center operations, saving money on IT staff and infrastructure.  *Example:* A startup can launch a global app without spending millions on servers.   **2. Scalability and Flexibility**  * Cloud resources can **instantly scale up or down** to match business needs. * Seasonal companies (like e-commerce during holidays) can handle traffic spikes easily.  *Example:* Amazon scales its cloud services every Black Friday to handle massive online sales traffic.  **3. Supports Remote and Global Teams**  * Employees can securely access files, apps, and systems from anywhere. * Encourages collaboration through tools like Microsoft Teams, Google Workspace, and Slack. *Example:* A company with offices in different countries can collaborate on the same cloud-based documents in real time.  **4. Access to Advanced Technologies**  * Cloud platforms offer AI, machine learning, and analytics tools without needing in-house expertise. * Businesses can use cloud-based analytics to gain insights from customer data and improve decision-making.  Example: Retailers use cloud AI to forecast demand and personalise marketing.  **5. Enhanced Security and Compliance**  * Top providers (AWS, Azure, Google Cloud) invest heavily in cybersecurity and compliance certifications. * Data encryption, identity management, and backup systems protect against loss and breaches.  *Example:* Financial institutions use secure cloud environments to meet strict data regulations.  **6. Faster Innovation and Time-to-Market**  * Cloud platforms allow developers to quickly build, test, and deploy applications. * This agility helps businesses experiment with new ideas and respond to market changes faster. *Example:* A software company can roll out updates weekly instead of quarterly  **7. Business Continuity and Disaster Recovery**  * Cloud backups protect critical data from physical disasters (fires, floods, etc.). * Rapid data recovery minimizes downtime and loss. *Example:* If a company’s local servers fail, they can instantly restore systems from the cloud.  **8. Sustainability and Efficiency**  * Cloud providers use large, optimized data centers that are more energy-efficient than most on-premises setups. * Helps companies meet sustainability goals.  **Real-World Impact Example** **Netflix** runs almost entirely on AWS. The cloud allows it to:   * Stream to millions of users worldwide, 24/7. * Scale resources automatically during peak hours. * Deploy updates and new features continuously.   Without the cloud, that level of global service and reliability would be nearly impossible. |
| What’s the alternative to cloud computing? | The main alternatives to cloud computing are:   1. **On-Premises Infrastructure:**  In this model, an organization owns and maintains its own servers, storage, and networking equipment. It offers full control over data and systems but involves high setup and maintenance costs and limited scalability. 2. **Colocation:**  The organization owns its hardware but rents space in a third-party data center that provides power, cooling, and physical security. It reduces infrastructure management burden but still requires the business to maintain its equipment. 3. **Private Cloud:**  A private cloud provides cloud-like features such as virtualization and automation but is dedicated to a single organization. It offers greater control and security but is expensive to set up and maintain. 4. **Hybrid Cloud:**  This combines on-premises systems with public cloud services. It allows sensitive data to remain in-house while less critical workloads run in the cloud. However, it is more complex to manage. 5. **Edge Computing:**  In edge computing, data is processed closer to the source of generation (e.g., IoT devices) rather than in a central cloud. It reduces latency and improves performance but has limited scalability. |
| What cloud providers can we use, what are their features and functions? | |  |  |  |  | | --- | --- | --- | --- | | Features | AWS | Azure | GCP | | Provider Overview | Launched by Amazon in 2006; the largest and most mature cloud provider with a broad range of services. | Launched by Microsoft in 2010; deeply integrated with Windows Server, Active Directory, and Microsoft 365. | Launched by Google in 2011; strong in analytics, AI, and open-source technologies. | | Compute | EC2, Lambda (serverless). | Virtual Machines, Azure Functions. | Compute Engine, App Engine, Cloud Functions. | | Storage | S3, EBS, Glacier. | Blob Storage, Disk Storage, Azure Backup. | Cloud Storage, Persistent Disk, Nearline/Coldline. | | Databases | RDS, DynamoDB, Redshift. | SQL Database, Cosmos DB, Synapse. | Cloud SQL, Firestore, BigQuery. | | Networking | VPC, Route 53, CloudFront. | Virtual Network, Azure CDN. | VPC, Cloud CDN, Cloud DNS. | | AI & ML | SageMaker, Rekognition. | Azure AI, Cognitive Services. | Vertex AI, AutoML, AI APIs. | | Containers | ECS, EKS (Kubernetes). | AKS (Kubernetes). | GKE (Kubernetes). | | Security | IAM, Shield, GuardDuty. | Active Directory, Sentinel. | Cloud IAM, BeyondCorp. | | Pricing | Pay-as-you-go, spot/reserved options. | Pay-as-you-go, hybrid benefits. | Pay-as-you-go, sustained-use discounts. | | Global Reach | Widest coverage worldwide. | Large enterprise-focused network. | Growing presence in key regions. | | Typical Use Case | Scalable web apps (e.g., Netflix, Airbnb). | Enterprise systems (e.g., HSBC, Adobe). | Data and AI workloads (e.g., Spotify, Snapchat). |   **In summary:**   * **AWS:** Best for scalability and variety of services. * **Azure:** Best for enterprises using Microsoft products. * **GCP:** Best for AI, analytics, and developer-focused environments. |

# Day 1: Task 2

Please research the below cloud offerings, explain what they are and examples of use cases.

|  |  |  |
| --- | --- | --- |
| Cloud Offerings | Explain what it is | When / how might you use this service in the real-world? |
| IaaS (Infrastructure as a service) | Provides virtualised computing resources over the internet, such as servers, storage, and networking. Users manage operating systems and applications while the provider handles the physical infrastructure. | Used when organisations want full control over their IT setup without owning physical servers — for example, hosting websites, developing applications, or running enterprise systems.  Example**:** AWS EC2, Microsoft Azure Virtual Machines, Google Compute Engine. |
| PaaS (Platform as a service) | Offers a cloud-based platform that includes infrastructure, operating systems, and development tools to build, test, and deploy applications. The provider manages everything except the application itself. | Used by developers to create and deploy apps quickly without managing servers — ideal for software development, API creation, or mobile app backends.  Example: Google App Engine, AWS Elastic Beanstalk, Microsoft Azure App Service. |
| SaaS (Software as a service) | Delivers fully functional software applications over the internet. Users access the software through a web browser; the provider manages everything including infrastructure, platforms, and updates. | Used for everyday business operations such as email, file storage, CRM, or collaboration tools.  Example**:** Google Workspace, Microsoft 365, Salesforce, Zoom. |

# Day 1: Task 3

Please research the below terms and explain what they are, when they would be appropriate and a real-world example of where it could be implemented (i.e. what type of organisation).

|  |  |
| --- | --- |
| Public Cloud | Cloud services offered over the internet by third-party providers, shared among multiple customers. Ideal for businesses needing scalable resources without heavy upfront investment.Real-world example: E-commerce platforms like Shopify or Netflix, which require scalable infrastructure to handle varying traffic loads. |
| Private Cloud | Cloud infrastructure used exclusively by a single organisation. Suitable for companies with stringent security, compliance, or performance requirements.  Real-world example: Financial institutions such as banks or insurance companies, which handle sensitive customer data and require high security. |
| Hybrid Cloud | A combination of public and private clouds, allowing data and applications to be shared between them. Useful for businesses that need to balance scalability with control over sensitive data.  Real-world example: Healthcare providers like hospitals, which store patient data privately but may use public cloud services for less sensitive applications. |
| Community Cloud | Shared cloud infrastructure for a specific community of organisations with common concerns, such as compliance or security. Appropriate for industries with shared regulatory requirements.  Real-world example: Government agencies or educational institutions collaborating on research projects, sharing resources while meeting compliance standards. |

# Day 2: Task 1

Describe, with examples, the **three** major areas that the Computer Misuse Act deals with.

|  |  |  |
| --- | --- | --- |
| Area | Description | Example |
| Unauthorized Access to Computer Material | Gaining access to any computer system or data without permission. | A hacker breaking into a company’s database to view confidential customer information. |
| Unauthorized Access with Intent to Commit or Facilitate a Crime | Accessing a computer system without permission **with the intention of committing further criminal activity**, such as fraud or theft. | Using stolen login credentials to transfer money from a company’s online bank account. |
| Unauthorized Modification of Computer Material | Changing, deleting, or adding data or programs without permission, including introducing viruses or malware. | Deploying ransomware on an organization’s network to encrypt files and demand payment. |

The computer misuse act 1990 is an act where an individual can be criminalised because of computer related offense. Describe three extra powers that the Police and Justice Act 2006 (Computer Misuse) has added.

|  |
| --- |
| Description |
| Making, supplying, or obtaining articles for computer misuse- It is now an offense to create, sell, or obtain software, programs, or tools intended for hacking or committing computer crimes.  Example: Creating or distributing password-cracking software or malware for others to use. |
| Increased penalties for serious computer misuse - The Act increased maximum sentences for certain offenses, making them more severe for serious or repeated crimes.  Example: A hacker causing significant financial loss to a company could face longer imprisonment than under the 1990 Act alone. |
| Offenses related to denial-of-service attacks - Specific criminalization of attacks that disrupt services, such as Distributed Denial-of-Service (DDoS) attacks.  Example: Launching a DDoS attack against a bank’s website, making it inaccessible to customers. |

Look at the below website to answer the questions:

<https://www.gov.uk/personal-data-my-employer-can-keep-about-me>

|  |
| --- |
| Write down three items of data which a company can store about an employee. |
| **Name** |
| **Address** |
| **Employment history with the organisation** |

|  |
| --- |
| Give three more examples of data that an employer can only store if they first get the employee’s permission. |
| Health and medical conditions |
| Religion |
| Trade union membership |

Conduct further research to answer the below questions.

|  |  |
| --- | --- |
| Question | Answer |
| |  | | --- | |  |   Shein vs. Nelly: A Swedish court ruled that Shein's subsidiary used Nelly's copyrighted photos without permission on Shein's Swedish website. | Shein vs. Nelly: A Swedish court ruled that Shein's subsidiary used Nelly's copyrighted photos without permission on Shein's Swedish website. |
| Provide one example of: Plagiarism | Melania Trump's 2016 RNC Speech: Portions of her speech were found to be similar to Michelle Obama's 2008 Democratic National Convention speech, leading to accusations of plagiarism. |
| What are two consequences of copyright infringement and software piracy? | * Legal Penalties: Individuals and organizations can face fines and imprisonment. For example, in the UK, copyright infringement can lead to criminal charges under the Copyright, Designs and Patents Act 1988. * Reputational Damage: Companies found guilty of piracy or infringement may suffer loss of consumer trust and damage to brand image. |
| Give three possible consequences for individuals when using pirated software | * Legal Risks: Using pirated software is illegal and can lead to fines or legal action. * Security Threats: Pirated software often contains malware, putting personal data and systems at risk. * Lack of Support: Users cannot access official updates or customer support, potentially leaving systems vulnerable. |

Listed below are some laws which we have covered today:

1. Computer Misuse Act 1990

2. Police and Justice Act 2006 (Computer Misuse)

3. Copyright, Designs and Patents Act 1988

4. Copyright (Computer Programs) Regulations 1992

5. The Health and Safety (Display Screen Equipment) Regulations 1992

6. Data Protection Act 2018

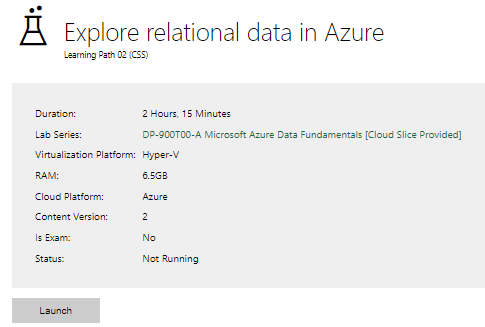
7. Consumer Rights Act 2015

* Insert a number in the first column of each row to match each of the statements with one of the above Acts.
* One of statements is incorrect and not illegal. For this statement, write ‘Not illegal’.

|  |  |
| --- | --- |
| **Act number** | **Clause** |
| 4 | With some exceptions, it is illegal to use unlicensed software |
| 7 | Any product, digital or otherwise, must be fit for the purpose it is supplied for |
| 1 | Unauthorised modification of computer material is illegal |
| 2 | It is illegal to create or use a hacking tool for penetration testing |
| 6 | Personal data may only be used for specified, explicit purposes |
| 5 | Employers must provide their computer users with adequate health and safety training for any workstation they work at |
| 2 | It is illegal to distribute hacking tools for criminal purposes |
| 3 | It is illegal to distribute an illicit recording |
| 6 | Personal data may not be kept longer than necessary |
| 1 | Gaining unauthorised access to a computer system is illegal |
| Not illegal | Employers must ensure that employees take regular and adequate breaks from looking at their screens |
| 2 | It is illegal to prevent or hinder access (e.g. by a denial-of-service attack) to any program or data held in any computer |
| 6 | Personal data must be accurate and where necessary kept up to date |

# Day 3: Task 1

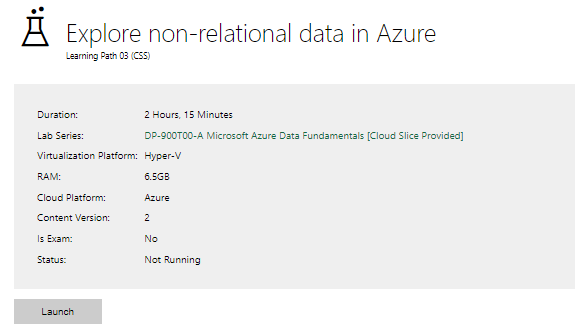
Please complete the below lab (3) *‘Explore relational data in Azure’* and paste evidence of the completed lab in the box provided.



|  |  |
| --- | --- |
| Completed lab |  |

# Day 3: Task 2

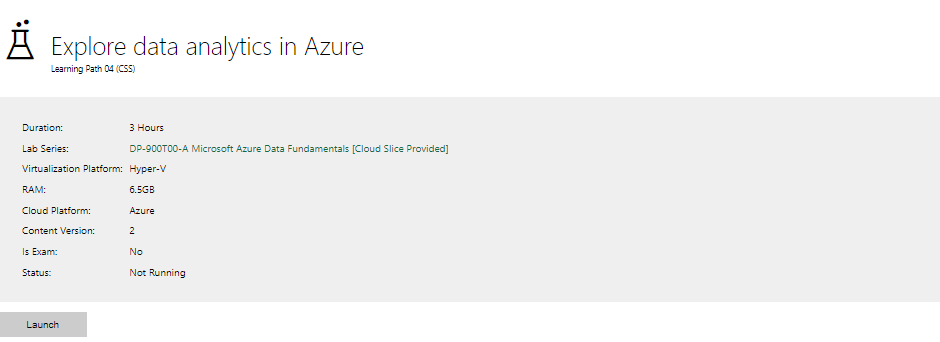
Please complete the below lab (4) *‘Explore non-relational data in Azure’* and paste evidence of the completed lab in the box provided.



|  |  |
| --- | --- |
| Completed lab |  |

# Day 3: Task 3

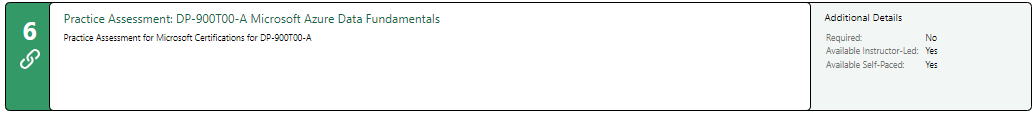
Please complete the below lab (5) ‘Explore data analytics in Azure’ and paste evidence of the completed lab in the box provided.



|  |  |
| --- | --- |
| Completed lab |  |

# Day 4: Task 1

In your teams, complete the Azure DP-900 practice exam and paste your result below – this is open book and please research and discuss your answers as a team.



|  |  |
| --- | --- |
| Result |  |

# Day 4: Task 2

#### **1. Scenario Background**

"Paws & Whiskers" is a growing pet shop that aims to improve its business by analysing sales, customer information, and inventory data. Currently, the data is collected manually or stored in spreadsheets. Management is interested in transitioning to Microsoft Azure to streamline data storage, analysis, and reporting, enabling them to make data-driven decisions.

#### **2. Data Laws and Regulations**

Identify and explain the data laws and regulations relevant to handling customer data within the proposal. Ensure you cover the following points:

* **GDPR Compliance**: Highlight the importance of adhering to the General Data Protection Regulation (GDPR), particularly as it relates to storing and processing customer information.
* **Data Protection Act (DPA) 2018**: Outline how the DPA 2018 may affect the way "Paws & Whiskers" collects and stores data, ensuring compliance with UK laws on data privacy.
* **Other Industry Standards**: Research any additional data protection standards or regulations that may apply to pet shop data, particularly if they involve sensitive or payment information.

#### **3. Azure Service Recommendations**

Recommend Microsoft Azure services that would suit the company’s data analysis needs and explain why these services are suitable. Your recommendations should include:

* **Data Storage**: Identify suitable storage options, such as **Azure Blob Storage** or **Azure SQL Database**, and discuss the benefits of each for storing large datasets, including inventory, sales transactions, and customer details.
* **Data Analysis Tools**: Recommend tools such as **Azure Machine Learning** for customer behaviour analysis or **Azure Synapse Analytics** for analysing sales trends.
* **Data Integration and Automation**: Explain how services like **Azure Data Factory** could automate data collection and integration processes, improving efficiency.

#### **4. Data Types and Data Modelling**

Define the types of data "Paws & Whiskers" will need to work with and describe your approach to data modelling:

* **Data Categories**: Identify key data types, such as customer demographics, transaction history, pet inventory, and product categories.
* **Data Modelling Approach**: Outline how you would structure this data using a relational model or a data warehouse approach, considering factors like tables, entities, relationships, and primary keys.

#### **5. Data Storage Formats and Structures in Azure**

Discuss how you would store data within Azure and the formats you would recommend:

* **Data Formats**: Specify recommended formats (e.g., CSV for raw data imports, JSON for structured data, Parquet for analytics) and explain why these formats are suitable for specific data types.
* **Data Security and Encryption**: Include recommendations for securing data using Azure’s built-in encryption features and access controls to ensure compliance with data privacy regulations.

#### **6. Additional Considerations**

Provide any other considerations that might enhance data handling and efficiency in Azure, such as:

* **Backup and Disaster Recovery**: Outline a backup plan using **Azure Backup** or **Azure Site Recovery** to safeguard against data loss.
* **Data Visualisation**: Discuss potential use of **Power BI** within Azure for creating dashboards that provide management with real-time insights into sales and customer trends.
* **Future Scalability**: Comment on how Azure services can scale as the business grows, accommodating larger datasets and more complex analyses.

### **Submission Guidelines:**

1. **Structure**: Ensure your report is well-organised, with sections for each task (e.g., Data Laws, Azure Services, Data Types, etc.).
2. **Formatting**: Include headings, bullet points where appropriate, and any visuals or diagrams that support your explanations.
3. **References**: Cite any resources or regulations referenced in the report.
4. **Length**: Aim for 1500-2000 words.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Whiskers Using Microsoft Azure****1. Scenario Background** *Paws & Whiskers* is a growing pet shop seeking to enhance business operations and decision-making by leveraging data analytics. Currently, the company relies heavily on **manual data collection and spreadsheets** to track sales, customer details, and inventory. This approach is **error-prone, time-consuming, and inefficient**, often leading to delays in understanding sales trends or inventory shortages. The management team has recognized the need for a **scalable, secure, and integrated data platform** to support informed business decisions.  By transitioning to **Microsoft Azure**, *Paws & Whiskers* aims to:   * **Centralize data storage** to consolidate disparate spreadsheets and manual records. * **Automate data collection and integration** across sales, inventory, and customer management systems. * **Perform advanced analytics** to identify purchasing trends, optimize inventory levels, and understand customer behavior. * **Generate real-time dashboards and reports** to enable rapid business decision-making. * **Scale infrastructure** to accommodate growth without disruption.   The implementation of Azure will enable *Paws & Whiskers* to **streamline business operations**, reduce errors, and increase operational efficiency. It provides a secure and compliant environment for storing sensitive customer data while supporting advanced analytical capabilities to improve overall business performance. **2. Data Laws and Regulations** Compliance with data protection laws is crucial when handling personal data. *Paws & Whiskers* must adhere to **GDPR**, the **UK Data Protection Act (DPA) 2018**, and other relevant standards to ensure the lawful and secure processing of customer data. **2.1 GDPR Compliance** The **General Data Protection Regulation (GDPR)** applies to organizations processing personal data of EU and UK residents. Even post-Brexit, the UK GDPR maintains similar standards. GDPR ensures that customer data is **collected, stored, and processed in a lawful, transparent, and secure manner**.  **Key GDPR Requirements:**   * **Lawful Processing:** Organizations must have a valid legal basis to process personal data. For *Paws & Whiskers*, this includes obtaining explicit consent from customers or fulfilling contractual obligations. * **Data Minimization:** Collect only the data necessary to perform business functions. * **Transparency:** Customers must be informed about what data is collected, why, and how it will be used. * **Data Subject Rights:** Customers have rights to access, correct, and request deletion of their personal data. * **Data Protection by Design and Default:** Implement security measures, including encryption, secure storage, and controlled access. * **Breach Notification:** Data breaches must be reported to regulatory authorities within 72 hours, and affected individuals must be informed when there is a high risk to their rights.   **Implications for Azure Implementation:**   * Store sensitive customer information in **Azure SQL Database** or **Azure Data Lake Storage** with **AES-256 encryption at rest**. * Implement **Role-Based Access Control (RBAC)** through Azure Active Directory to restrict access. * Maintain **audit logs** using Azure Monitor to track data access and changes. * Apply **network security rules** and secure connections (TLS/SSL) to protect data in transit.  **2.2 UK Data Protection Act 2018 (DPA 2018)** The **DPA 2018** complements the GDPR with UK-specific provisions and governs **how personal data is collected, stored, and processed** within the UK. It also outlines requirements for lawful data handling, record keeping, and accountability.  **Key Considerations:**   * **Lawful Processing:** Personal data must be collected and used fairly, transparently, and only for legitimate business purposes. * **Data Security:** Adequate security measures must be implemented to protect personal data from unauthorized access. * **Retention Policies:** Personal data should be retained only as long as necessary. * **Accountability:** Organizations must demonstrate compliance with data protection principles and maintain processing records.   **Azure Implementation Strategies:**   * Configure **data retention policies** in Azure SQL Database and Data Lake Storage. * Enable **auditing and monitoring** to track user access. * Implement **secure backup procedures** using Azure Backup and Site Recovery to ensure data integrity in case of accidental loss or disaster.  **2.3 Other Industry Standards**  * **PCI DSS Compliance:** If *Paws & Whiskers* processes credit card payments, it must comply with PCI DSS, which mandates encryption, secure networks, and access controls for payment data. * **ISO/IEC 27001:** Azure services comply with this international standard for information security, ensuring secure storage, transmission, and processing of sensitive data. * **Sensitive Data Handling:** Certain customer data, such as health-related pet information or payment details, requires heightened security measures, including tokenization, encryption, and restricted access.   By adhering to these regulations, *Paws & Whiskers* ensures **legal compliance, customer trust, and secure management of sensitive data**. **3. Azure Service Recommendations** Selecting the right Azure services is critical to achieving **efficient data storage, analysis, and reporting**. **3.1 Data Storage Solutions**  |  |  |  | | --- | --- | --- | | **Service** | **Use Case** | **Benefits** | | Azure SQL Database | Customer information, sales transactions | Fully managed, relational structure, supports ACID transactions, automatic backups, scalable | | Azure Blob Storage | Raw inventory data, CSV files, JSON files | Cost-effective for unstructured and semi-structured data, easily integrated with analytics services | | Azure Data Lake Storage Gen2 | Large analytics datasets | Optimized for big data, hierarchical namespace, parallel processing, integrates with Azure Synapse and Databricks |   **Rationale:** Azure SQL Database is ideal for structured data requiring relational integrity, while Blob and Data Lake Storage are suitable for large, unstructured datasets like historical sales logs and inventory files. **3.2 Data Analysis Tools**  * **Azure Synapse Analytics:** Enables integration of structured and semi-structured data, supports large-scale analytics, and allows real-time insights into sales trends and inventory patterns. * **Azure Machine Learning:** Supports predictive analytics for customer behavior, sales forecasting, and inventory optimization. * **Power BI:** Provides interactive dashboards, real-time visualization, and reporting capabilities. Directly integrates with Azure SQL Database and Synapse Analytics for live data querying.  **3.3 Data Integration and Automation**  * **Azure Data Factory (ADF):** Orchestrates data pipelines, automates extraction from spreadsheets, POS systems, and online platforms, transforms data, and loads it into centralized storage. * **Benefits:** Reduces manual errors, ensures consistent data, enables scheduled batch processing, and supports scalable ETL processes.  **4. Data Types and Data Modelling****4.1 Data Categories**  * **Customer Data:** Names, addresses, emails, phone numbers, purchase history. * **Sales Transactions:** Transaction ID, products, quantities, prices, date/time, payment method. * **Inventory Data:** Stock levels, product types, supplier details, reorder dates. * **Product Categories:** Pet food, accessories, health products, toys.  **4.2 Data Modelling Approach**  * **Relational Database Model:**   + **Tables:** Customers, Transactions, Products, Inventory.   + **Primary Keys:** CustomerID, TransactionID, ProductID.   + **Foreign Keys:** Link transactions to customers and products. * **Data Warehouse Approach:**   + **Star Schema:**     - **Fact Table:** SalesTransactions (contains measurable metrics like revenue, quantity).     - **Dimension Tables:** Customers, Products, Time, Store.   + Supports **aggregated reporting**, trend analysis, and advanced analytics.   **Rationale:** A combination of relational databases for operational tasks and a data warehouse for analytics provides **efficiency, integrity, and scalability**. **5. Data Storage Formats and Structures in Azure****5.1 Recommended Data Formats**  |  |  |  | | --- | --- | --- | | **Format** | **Use Case** | **Reason** | |  |  |  | | CSV | Spreadsheet imports, raw data | Simple, universally supported, ideal for batch imports | | JSON | Customer and semi-structured data | Flexible, easy to process with Azure tools | | Parquet | Analytical datasets | Columnar format optimized for query performance and storage efficiency |  **5.2 Data Security and Encryption**  * **Encryption at Rest:** Azure SQL Database and Data Lake Storage automatically encrypt data using **AES-256**. * **Encryption in Transit:** Use **TLS/SSL** for secure data transmission. * **Access Control:** Implement **RBAC** and **Azure AD authentication**. * **Auditing and Monitoring:** Enable **Azure Monitor** to track access, changes, and potential breaches.  **6. Additional Considerations****6.1 Backup and Disaster Recovery**  * **Azure Backup:** Schedule automated backups for databases and storage. * **Azure Site Recovery:** Provides disaster recovery capabilities in case of service outages or hardware failures. * **Recovery Strategy:** Daily incremental backups, weekly full backups, and offsite replication ensure data integrity.  **6.2 Data Visualization**  * **Power BI Dashboards:** Real-time visualizations of sales, inventory, and customer trends. * **KPIs:** Top-selling products, monthly revenue trends, stock shortages, and customer acquisition metrics. * **Integration:** Connects to Azure SQL Database or Synapse Analytics for live queries.  **6.3 Future Scalability**  * Azure services offer **elastic scaling**:   + **Azure SQL Database** scales compute and storage on-demand.   + **Data Lake Storage Gen2** |

|  |
| --- |
| **Course Notes** |

It is recommended to take notes from the course, use the space below to do so, or use the revision guide shared with the class:

|  |
| --- |
| **Additional Information** |

We have included a range of additional links to further resources and information that you may find useful, these can be found within your revision guide.

**END OF WORKBOOK**

**Please check through your work thoroughly before submitting and update the table of contents if required.**

**Please send your completed work booklet to your trainer.**